EOSDIS Core System

Architecture Overview

Steve Fox

Roadmap

Origin of Current ECS Architecture

Key Architecture Concepts

ECS Capabilities and Interfaces of Interest to ESIS

Example Scenarios

Some History

GCDIS / UserDIS Study

- **NRC:** *"Provision of common GCDIS (and UserDIS) software, database structures, and technical infrastructure for an interoperable network"*

- **Results of Study:** There are components of GCDIS/userDIS which ECS can provide without leaving its mission envelope and without a lot of additional cost, by carefully choosing the appropriate architectural direction.

Boundaries

- **Cannot Compromise EOS Primary Mission Objectives**
  - **Must Preserve EOS Data (An Extremely Valuable National Resource)**
  - **Cannot Endanger The High-Volume Data Ingest and Production Operation**
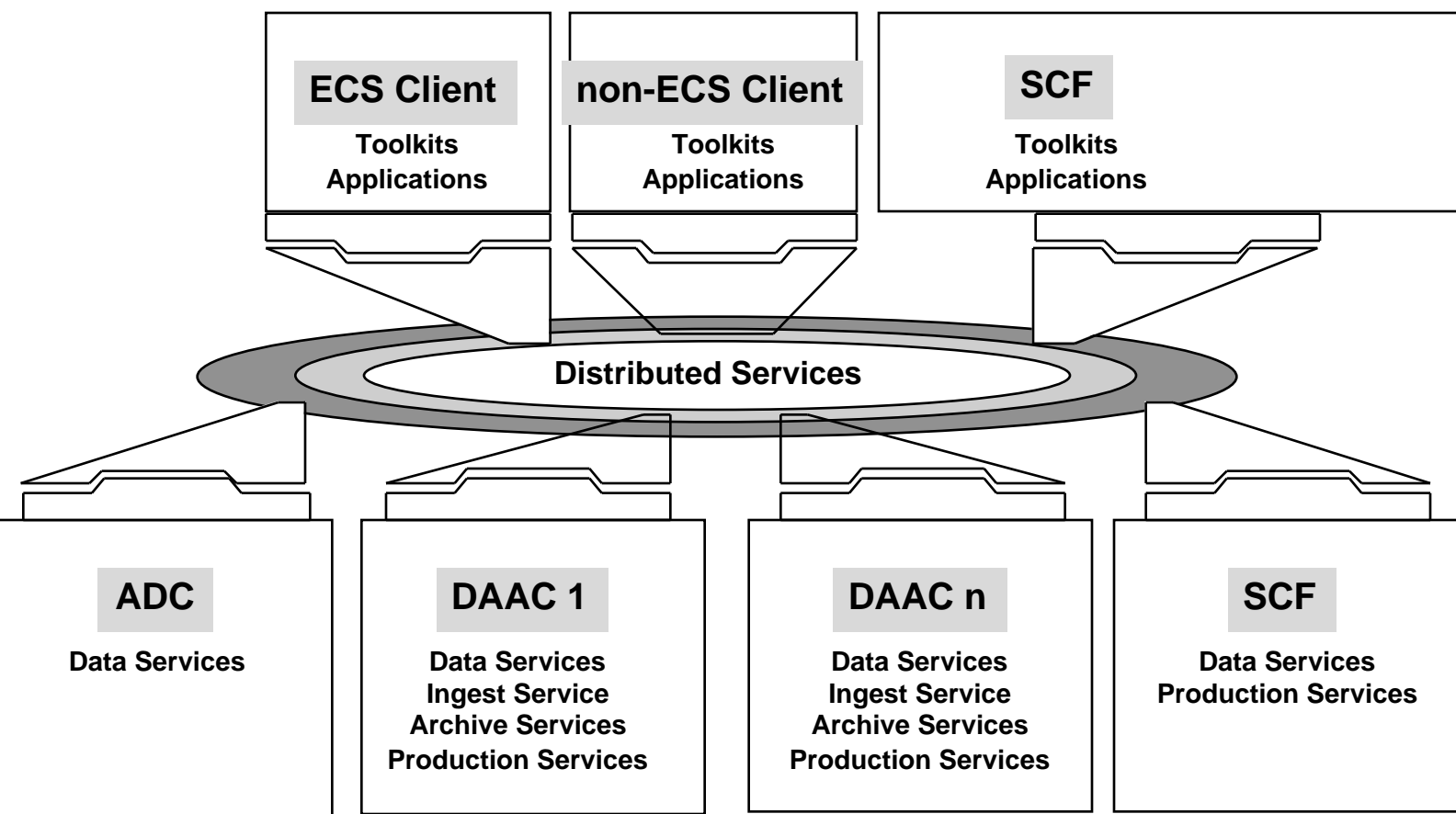 - **Cannot Degrade Critical Capabilities of EOS Researchers**

Key Architecture Concepts

Service Distribution & Site Autonomy

ECS Client	non-ECS Client	SCF
Toolkits **Applications**	**Toolkits** **Applications**	**Toolkits** **Applications**

Distributed Services

ADC	DAAC 1	DAAC n	SCF
Data Services	**Data Services** **Ingest Service** **Archive Services** **Production Services**	**Data Services** **Ingest Service** **Archive Services** **Production Services**	**Data Services** **Production Services**

ECS Program Schedule

993	1994	1995	1996	1997	1998	1999	2000	2001	2002

▲ Award of Contract

TRMM ▲   AM-1 ▲   ▲ ADEOS II   End of Contract ▲

LANDSAT 7 ▲   ▲ COLOR   ▲ PM-1

▲▲▲ ▲▲ ▲▲ ▲▲▲ Evaluation Packages and Toolkits

TRMM Mission (Release A)

PDR   CDR

Science Environment (IR-1)

AM/Landsat 7 (Release B)
ADEOS II

PM/SAGE III
ACRIM/R ALT (Release C)

CHEM/AL
(Release

DAAC Operations DAAC Activation   Full Operations

Mission Operations EOC Activation   EOC Operations

ECS Architecture Concepts

ECS Production and Archive functions have characteristics which are important for ESIS

Production

- Designed to allow routine and on-demand production
- Site production is autonomous
- Inter-site coordination policy supported by tools, not hard-coded into design
- Production sized for producing standard products at keep-up rates & for reprocessing

Archiving

- Data = Data: no artificial distinction between "Directory", "Inventory", and "Granules"
- Data = Services: Data is accessed via "Services" -- no artificial distinction between data access services and other types of services available for data (e.g., browse, subset, and on-demand production)

ECS Architecture Concepts

ECS Capabilities are Integrated with Local System Management Infrastructure.
— System Management is not Centralized —

System Management

- **Local System Management - each site manages and operates its own resources and maintains its own schedules**

- **Local Service Management - each site monitors its service levels (e.g., performance, errors, complaints), and tracks its service requests**

- **Local User Support - each site has user help services**

- **Intersite System Monitoring - information forwarded by sites is consolidated for monitoring and reporting purposes**

ECS Architecture Concepts

ECS is built on top of a distributed object framework.

Interfaces are modeled as Object Frameworks

- Example: Session/Request
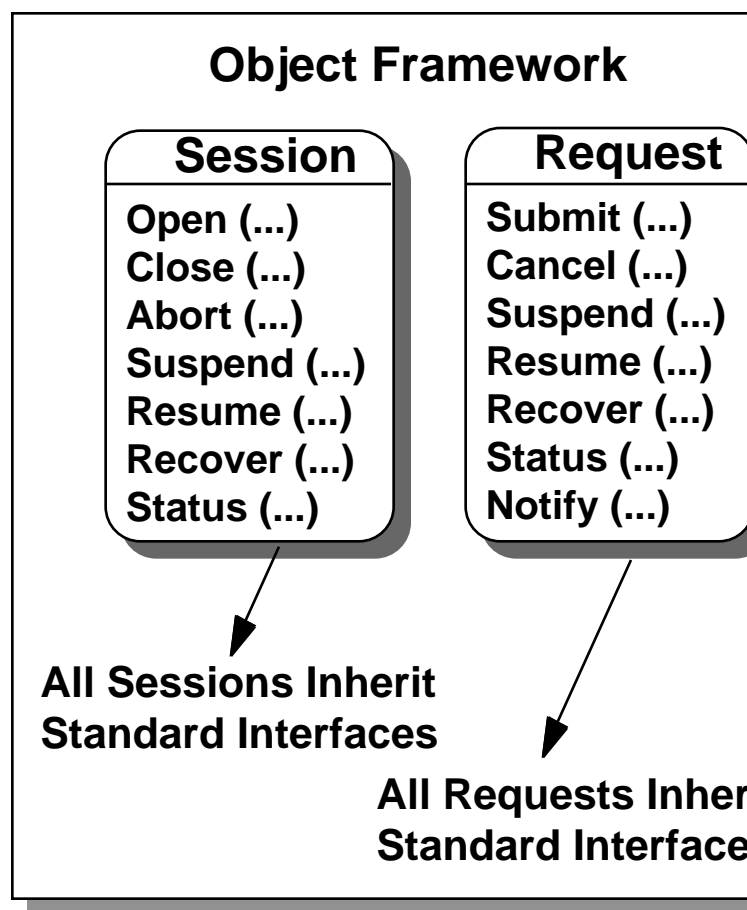- Framework defines interface standard
- Service providers and clients inherit that standard
- Distributed computing infrastructure provided by DCE

"Legacy" interfaces are used where distributed objects are not yet ready for prime time

- Example: bulk data transfers
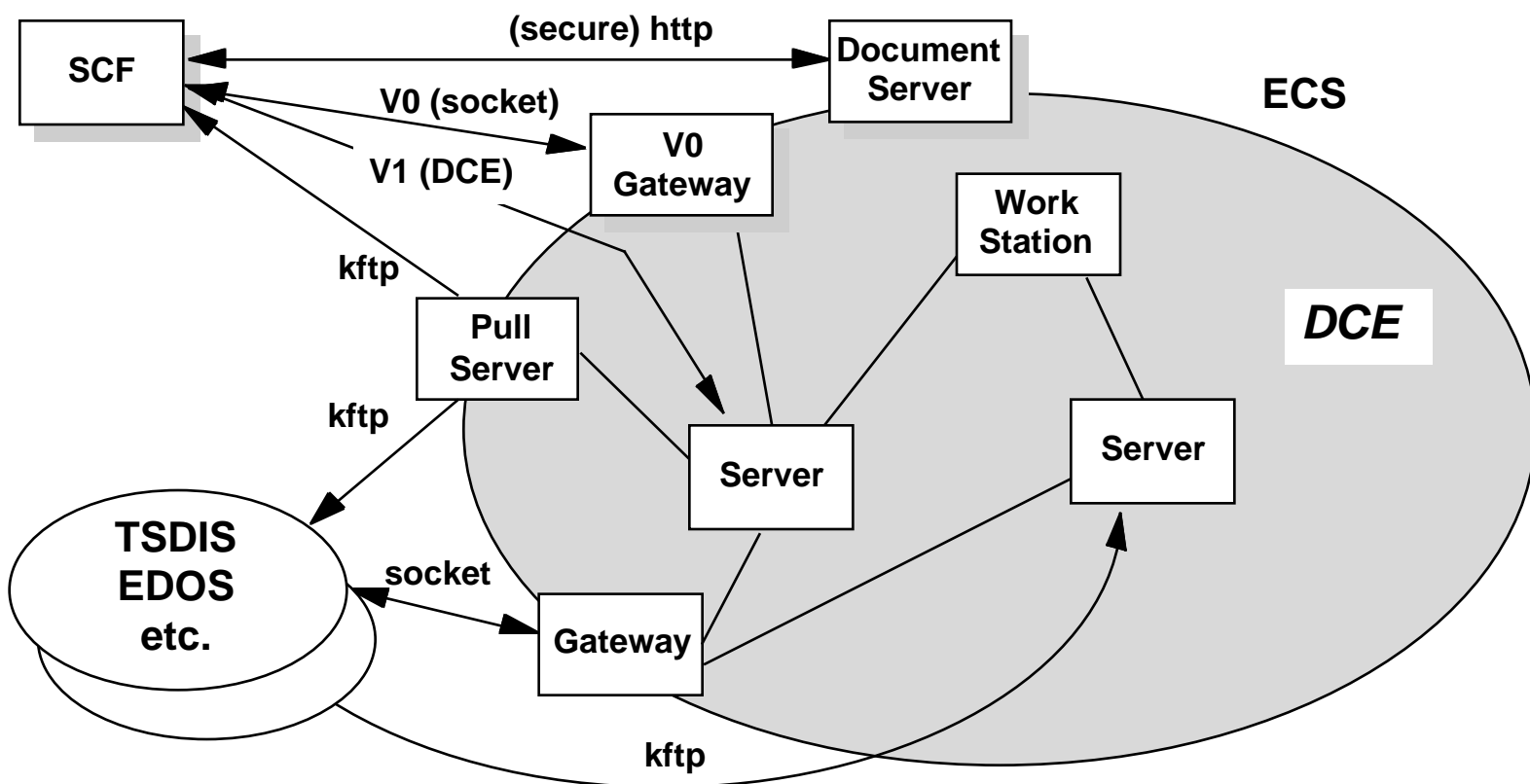- They are encapsulated with "object methods"

Object Framework

Session
Open (...)
Close (...)
Abort (...)
Suspend (...)
Resume (...)
Recover (...)
Status (...)

Request
Submit (...)
Cancel (...)
Suspend (...)
Resume (...)
Recover (...)
Status (...)
Notify (...)

All Sessions Inherit Standard Interfaces

All Requests Inherit Standard Interfaces

ECS Architecture Concepts

ECS uses gateways to interface with external entities
http/HTML, ftp & sockets

ECS Architecture Concepts

R
CNE
(GSFC Only)   R

SCF   R

R

EBnet   R

Internet

R

any
DAAC
MAN   R

R   R   R   R

isoCell --
*under
consideration*

Server   Server   Server

R

GSFC ECS LAN

R   R

GSFC Cell

EBnet

any DAAC Cell

R   R

any DAAC ECS LAN

Server   Server   Server
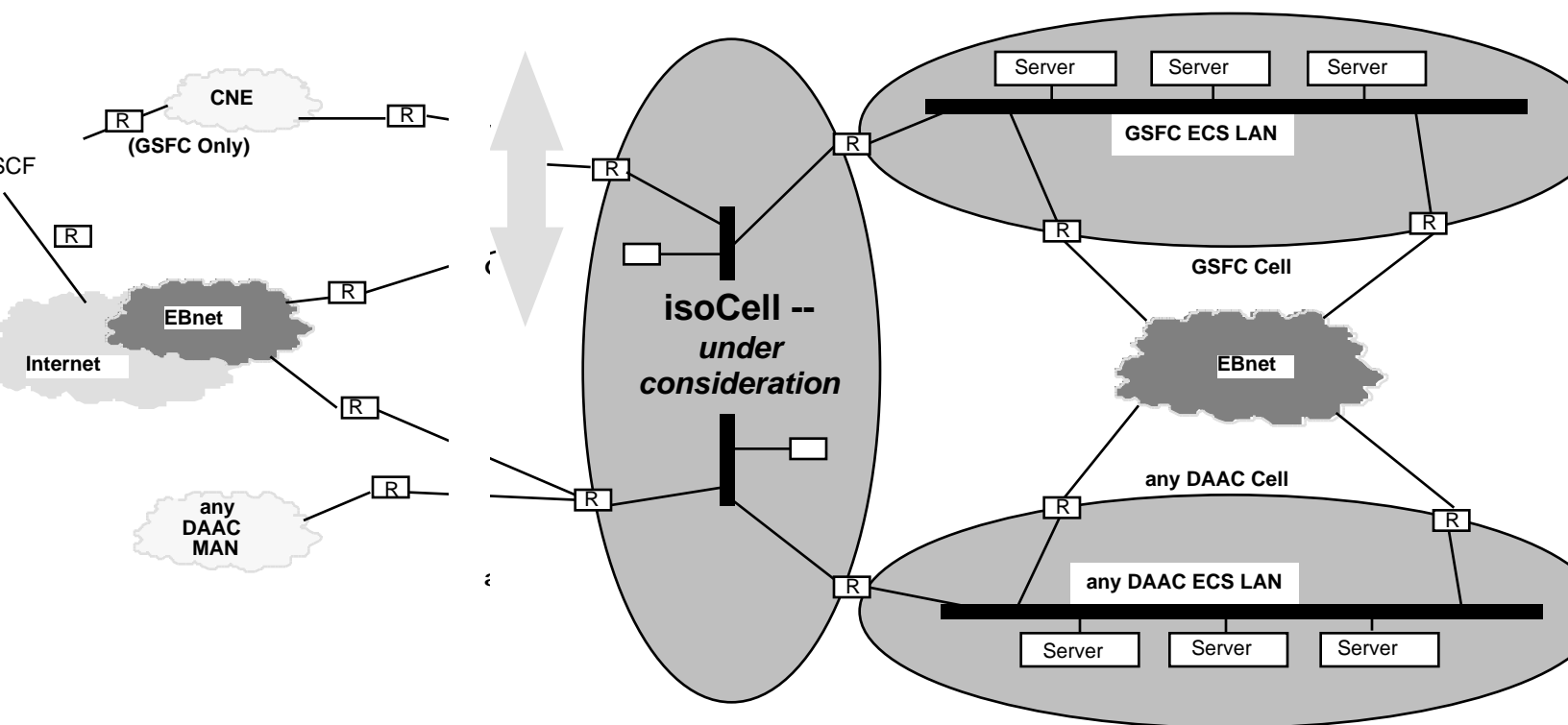
ECS relies on DCE for security - each site is a DCE cell

Local Security Management: each site manages is own security

Cross cell authentication and authorization

ECS Architecture Overview

ECS Interface Overview

Within ESIS, sites may have different roles and choose different levels of integration with the federation

- Roles and Levels of Integration are TBD

ECS DAAC sites have a very high level of integration into the "system"

The interfaces and capabilities offered by ECS which are presented in the following slides are organized around three roles

- Producers

- Archivers

- System Monitors

Interfaces Available for Producers

Supporting on-demand production orders
- *distributed object based i/f*
- accept request, provide status, provide notification of exceptions, etc.
- However, current i/f assumed to be with an "ordering agent" (Data Server is the ordering agent in ECS) which always functions as go-between

Subscribing to data; retrieving granules & meta data
- *distributed object based i/f*

Ingesting data
- *non-object interfaces (ftp, sockets, http)*
- (ingest only used for non-ECS data)

Storing data
- *distributed object based i/f*
- language and standards for providing meta-data

Interfaces Available for Producers (cont)

Integrating science software into production environment

- **toolkit library**

Some interfaces needed by producers, such as

- returning granules to requester
- subscriptions to produced data (e.g., for QA) or production problems
- satisfying downstream dependencies at other sites

are Data Server interfaces

- in ECS, a producer always has an archive

Interfaces Available for Archiver

Exporting advertisements and dictionary information
- *distributed object based i/f*

Accepting data products to be archived
- *distributed object based i/f*

Ingesting data
- *Internet gateways*
- (used for "non-ECS data")

Document search and access
- *Z39.50; http*

Data search and access
- *ESQL, distributed object based i/f; http gateway; Z39.50 gateway under consideration*

Accepting subscriptions
- *distributed object based i/f*

Distributing data & notifications
- *Internet gateways & distributed object based i/f*

Interfaces Available for System Monitors

Managing the site is the responsibility of the local site staff, but ECS will have a System Monitoring Capability (at GSFC).

Interfaces include

Collecting summarized management and fault information
- *file transfers, e-mail, distributed object based i/f*

Tracking requests across sites
- *HTTP/HTML or distributed object based i/f*

Coordinating schedules (e.g., ground events)
- *manual interfaces (assisted by tools)*

Data Standards

Core Science Data Model

- **follows FGDC**

Data Exchange Standard for Science Data

- **extension of HDF**

Document Standard

- **HTML (plus support for several proprietary formats)**

On-line Help Standard

- **still TBD**

ECS Capabilities for ESIS Providers

Data archive, data management, and distribution software
- **Includes significant COTS**

Science data production management and control software
- **Includes COTS**

Generic data ingest server

Implementations for interfaces to a system management infrastructure
- **based on SNMP & DCE**
- **(system management infrastructure itself is mostly COTS)**

Software to support data standards (e.g., HDF)

ECS Capabilities for ESIS Community

Advertising server

Data dictionary service

Distributed search agents (DIM) and local search gateway (LIM)

HTTP/HTML gateways

Z39.50 gateway under consideration

Loosely Coupled Provider

CASE

Provider has some interesting data and wants the rest of the communit[y] to know about them

Does not want to install a lot of new software to do that

May have the data accessible via an http server

- **Provider only uses advertising interface to announce his data**

Producer in Need
of Archive Capability

CASE

Provider already has a production system, but wants to install a data management component which readily links into the ESIS network and gives others access to his data

Provider

- installs ECS Data Server software
- interfaces his production with the front-end Data Server interfaces
- links existing DBMS and/or archive into the back-end interfaces of the Data Server
- may need to integrate any special access software ("user methods") into the Data Server (e.g., subsetting)
- goes through the necessary setup steps (schema, configuration data, access control files, network hookup)
- creates and submits advertisements

Value-Added Provider Needing A Production Environment

CASE

Value-added provider wants to make use of ESIS, e.g., get data which are of interest specifically to her user community and provide special products (e.g., tailored subsets, overlays from several sources, special formats).

Provider already has an archive and a set of existing interfaces via which users access the data, but needs a production system.

Provider

- installs ECS production control and ingest software and interfaces it with the existing archive (Data Server interfaces provide model)
- interfaces the science software with the ECS SDP toolkit as necessary
- writes any needed ingest conversion and validation routines
- coordinates production requirements & schedules with other sites
- goes through the necessary setup steps (subscribe to products needed from other ESIS sites, create production control data & product meta files, etc)
- creates and submits advertisements

Integrating Into ESIS

CASE

...ata provider wants to become part of the "One-Stop-Shopping" network...

...rovider

- **determines the dictionary / vocabulary and data names under which to "publis... the data**
- **creates the schema and any dictionary descriptions needed for publishing**
- **creates a local search agent (perhaps re-using ECS LIM software)**
- **goes through the necessary setup steps (e.g., schema and dictionary export, access control files, network hookup)**
- **creates and submits advertisements**

Summary

- **ECS architecture supports distributed production and data management services with a focus on site autonomy**

- **ECS intersite interfaces can serve as an architectural model for ESIS**